Application of Mathematical Tools for Metabolic Design of Microbial Ethanol Production

Vassily Hatzimanikatis,* Marcel Emmerling, Uwe Sauer, James E. Bailey

Institute of Biotechnology, ETH Zurich, CH-8093 Zurich, Switzerland; telephone: 411 633 31 70; fax: 411 633 10 51

Received 7 July 1997; accepted 30 July 1997

Abstract: Many attempts to engineer cellular metabolism have failed due to the complexity of cellular functions. Mathematical and computational methods are needed that can organize the available experimental information, and provide insight and guidance for successful metabolic engineering. Two such methods are reviewed here. Both methods employ a (log)linear kinetic model of metabolism that is constructed based on enzyme kinetics characteristics. The first method allows the description of the dynamic responses of metabolic systems subject to spatiotemporal variations in their parameters. The second method considers the product-oriented, constrained optimization of metabolic reaction networks using mixed-integer linear programming methods. The optimization framework is used in order to identify the combinations of the metabolic characteristics of the glycolytic enzymes from yeast and bacteria that will maximize ethanol production. The methods are also applied to the design of microbial ethanol production metabolism. The results of the calculations are in qualitative agreement with experimental data presented here. Experiments and calculations suggest that, in resting *Escherichia coli* cells, ethanol production and glucose uptake rates can be increased by 30% and 20%, respectively, by overexpression of a deregulated pyruvate kinase, while increase in phosphofructokinase expression levels has no effect on ethanol production and glucose uptake rates. © 1998 John Wiley & Sons, Inc. *Biotechnol Bioeng* 58: 154–161, 1998.

Keywords: central carbon pathways; metabolic optimization; ethanol production

INTRODUCTION

Current knowledge of biochemical systems is composed of a vast set of data that accumulates at an increasing rate. Advances in analytical methods and development of sophisticated instrumentation have provided the tools that allow us to know more than we can understand. Comprehension is complicated by the high complexity of living organisms. This complexity increases from unicellular organisms to multicellular structures, such as tissues and organs, to differentiated organisms.

Recombinant DNA and other modern genetic technology allow the introduction of precise changes in the genetic determinants of cell function and, thus, the engineering of metabolic activities for novel and/or improved functions. The application of molecular genetics and engineering tools for the useful manipulation of cellular processes is defined as *metabolic engineering* (Bailey, 1991). This definition identifies the engineering virtues of the manipulation of metabolism: putting together the available information, employing intuition, and tackling the problem. However, as in many engineering approaches, solution of the problem has not been always achieved. Moreover, the discovery of some of the most profitable biotechnological applications has been serendipitous.

Current approaches and methodologies in the biological sciences, while providing understanding of isolated cellular processes and subsystems, do not promote understanding of the simultaneous contributions of these subsystems to overall cellular metabolism. It is the organization of these subsystems, which are themselves relatively small but elaborate networks, in large complex networks which cause intuition and piece-wise knowledge to fail as tools for rational metabolic engineering (and, similarly, for accomplishing molecular medicine). Moreover, the high dimensionality of these networks offers an almost infinite number of alternative approaches toward the achievement of a goal. Most of the metabolic networks considered for engineered manipulation are composed of enzyme-catalyzed reactions with complicated stoichiometry, nonlinear kinetics, and superimposed regulatory structures. These regulatory structures include interactions of enzymes with compounds in the reaction networks, which often are not reactants or products of the reaction catalyzed by the enzymes in question, but serve to modulate the catalytic activity of the enzymes. Manipulation of a metabolic subsystem will change the metabolite concentrations and the metabolic activities of the subsystem. This manipulation and the resultant subsystem changes often drive responses in the subsystem's cellular environment that have not been taken into consideration.

The development of tools and frameworks that will organize the available biological knowledge and that will help in the analysis and design of metabolic networks is of im-

Correspondence to: J. E. Bailey

**Present address:* BioControl Group, DuPont Central Research Department, Experimental Station, Wilmington, DE 19880

Contract grant sponsors: Swiss Priority Program in Biotechnology; Swiss Federal Department of Science and Education

CCC 0006-3592/98/020154-08

mediate importance. These tools should be able to screen efficiently through an almost unlimited set of cellular modifications (realizable by genetic engineering technology), and report a small set of most promising options that can be further tested for their potential in the development of new products and improved bioprocesses.

MATHEMATICAL METHODS FOR ANALYSIS AND DESIGN OF METABOLIC SYSTEMS

Mathematical models of cellular processes have been used successfully in the past for the analysis of these processes, and they have provided useful initial directions for genetic improvements of the process of interest. The realization that mathematical models are the only way that net consequences of simultaneous, coupled, and often counteracting processes can be evaluated consistently and quantitatively has led to the growth of mathematical modeling in many biological and biotechnological areas. The analysis of these models requires development of associated mathematical and computational methods. Two mathematical and computational methods are proposed here as tools for obtaining useful guidance for experimental application of metabolic engineering.

One of the most common targets of metabolic engineering is the manipulation of metabolic functions, such as metabolic reaction rates and intracellular metabolite concentrations, in desired directions. Prediction of the changes in metabolic functions, which will occur after a change in any parameter affecting the metabolism, requires a kinetic representation relating fluxes to metabolite concentrations and metabolic parameters. In general, such models are not available, and the kinetic model must, therefore, be developed from measurements of reaction rates and metabolite concentrations. A variety of formulations are possible. A (log)linear kinetic model for metabolic reaction networks is presented here (Hatzimanikatis and Bailey, 1997a). The model does not require detailed information about the kinetic mechanisms of the reactions. It simply employs knowledge about the strength of interaction among the rates of enzyme-catalyzed reactions and the concentrations of the various metabolites, substrates, and regulators of the metabolic network. This kind of information can be derived from experimental and theoretical methods developed around metabolic control analysis (MCA), a sensitivity analysis framework that allows determination of quantitative indices such as control coefficients and elasticities (Cornish-Bowden and Cardenas, 1990; Fell, 1992; Kacser and Burns, 1973; Schlosser et al., 1993). Comparative studies between (log)linear models and nonlinear models based on common reaction mechanisms have been performed, demonstrating the satisfactory accuracy of (log)linear models in approximating the dynamic responses of metabolic networks to changes in metabolic parameters (Hatzimanikatis and Bailey, 1997a).

Unsuccessful attempts to engineer cellular metabolism by simply manipulating the amount of various enzymes suggest that engineering of the regulatory characteristics of the enzymes in a metabolic network offers greater potential for the achievement of desired metabolic properties. In many examples of small reaction networks with simple regulatory structures, changes in the regulation of one or two enzymes improved product formation significantly (Bailey, 1991; Katsumata and Ikeda, 1993). In large metabolic networks, or those closer to central carbon metabolism, the regulatory structures tend to be more elaborate. Moreover, the experimental difficulties in modifying regulatory interactions add another degree of complexity, because the common trial-and-error experimental approach is unfeasible. A production-oriented optimization framework has been recently developed (Hatzimanikatis et al., 1996). The framework considers the optimization of the performance of a metabolic pathway with respect to changes in the amounts of enzymes in the pathway and in modifications in the regulatory characteristics of those enzymes. The optimization study is undertaken using the kinetic description provided by the (log)linear kinetic model, used in this case to define constraints on the optimization in the form of steady-state mass balances of intracellular metabolites. Consideration of changes in enzyme regulation which abolish initially present inhibition or activation introduces binary decision variables into the optimization, resulting in a mixed-integer programming problem.

The development of these mathematical methods has been extensively presented elsewhere (Hatzimanikatis et al., 1996; Hatzimanikatis and Bailey, 1997a). We will present here an application to microbial ethanol production, and the results of this analysis will be compared to experimental data.

METABOLIC ENGINEERING OF MICROBIAL ETHANOL PRODUCTION

Ethanol production from sugar mixtures by microorganisms has been the subject of extensive research (Ingram and Doran, 1995). *Saccharomyces cerevisiae* and *Zymomonas mobilis* have been the most widely used organisms for ethanol production. However, in recent years *Escherichia coli* has been genetically engineered for ethanol production (Ohta et al., 1991). The main pathway involved in ethanol production is the glycolytic pathway; Figure 1 shows this pathway and its regulation in *S. cerevisiae*. The complexity of glycolysis and its regulation has been the main obstacle in many experimental attempts to increase ethanol production and, in general, to manipulate its functions by metabolic engineering (Boles et al., 1993; Schaaff et al., 1989). Mathematical modeling of the glycolytic pathway in various organisms and computer simulations have also been used for a better understanding of the performance of the pathway and its responses to genetic manipulations (Gallazo and Bailey, 1990; Hatzimanikatis and Bailey, 1997b; Heinrich et al., 1977; Joshi and Palsson, 1989; Schlosser et al., 1994).

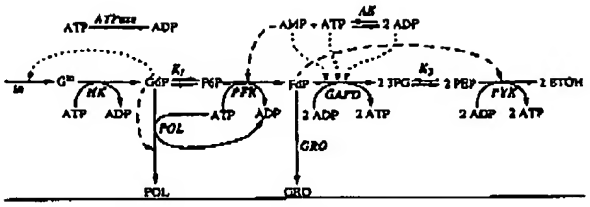

Figure 1. Anaerobic fermentation pathway of nongrowing yeast, *Saccharomyces cerevisiae*, with glucose as the sole carbon source. Chemical species: G^{in}, intracellular glucose; G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; FdP, fructose 1,6-diphosphate; 3PG, 3-phosphoglycerate; PEP, phosphoenolpyruvate; POL, polysaccharides; GRO, glycerol; ETOH, ethanol; ATP, adenosine triphosphate; ADP, adenosine diphosphate; AMP, adenosine monophosphate. Pathway steps and enzymes; in, glucose uptake; **HK**, hexokinase; K_1, equilibrium step; **PFK**, phosphofructokinase; **GAPD**, glyceraldehyde 3-phosphate dehydrogenase; K_2, equilibrium step; **PYK**, pyruvate kinase; **GRO**, glycerol production; **POL**, polysaccharide production; **ATPase**, net ATP consumption; **AK**, adenylate kinase. Solid arrows indicate reaction steps, dashed arrows indicate activation and dotted arrows indicate inhibition.

A (Log)linear Model of Microbial Ethanol Production

A mathematical model for the description of ethanol production by nongrowing *S. cerevisiae* has been constructed, and found to be consistent with experimental observations (Gallazo and Bailey, 1990; Schlosser et al., 1994). This model uses nonlinear rate expressions based on kinetic parameters estimated from in vitro studies, as well as expressions parametrized by in vivo perturbation experiments. The pathway features seven regulatory loops of activation and inhibition, and the kinetic expressions are strongly nonlinear. Based on this nonlinear model, we constructed a (log)linear model around a reference steady state. The corresponding parameters of the (log)linear model, i.e., the stoichiometric and elasticities matrices, are presented in the Appendix.

We performed a comparison between the (log)linear and the nonlinear models with respect to their dynamic responses to changes in metabolic parameters. Characteristic results for the responses of the two models to step and pulsed-periodic variations in glucose uptake are presented in Figure 2, and they demonstrate the excellent accuracy of the (log)linear model. The final steady-state differences for step responses are due to inherent limitations in MCA calculations of new steady states.

Excellent agreement between response characteristics of the (log)linear model and of the corresponding nonlinear model has also been observed for various complicated pathways (Hatzimanikatis and Bailey, 1997a). This agreement suggests that, in the absence of a nonlinear model, and when MCA quantities can be experimentally determined or estimated, the (log)linear model can be used as a good first approximation for the analysis of the dynamic response characteristics of metabolic systems.

Metabolic systems, like any nonlinear dynamic system, can exhibit a shift in time-averaged properties (such as metabolic fluxes and control coefficients) in response to fluctuations with zero-time-average value. Such fluctuations can arise, for example, in bioreactors when mixing is nonideal, a situation common in large-scale systems. Based on the excellent approximation provided by the (log)linear model, a method has been developed for the estimation of the performance of metabolic systems subject to periodic spatiotemporal variations of the system parameters and process operating conditions. Application of the method to the ethanol production pathway has shown that, for periodic fluctuations in extracellular glucose concentration, time-average flux control coefficients vary strongly, and not monotonically, with the period of the external fluctuations (Hatzimanikatis and Bailey, 1997a). These observations are particularly significant because they imply that decisions taken at a certain stage of process scale-up, with respect to which enzyme in a pathway should be genetically manipulated, could have no effect, or even be counterproductive at another scale.

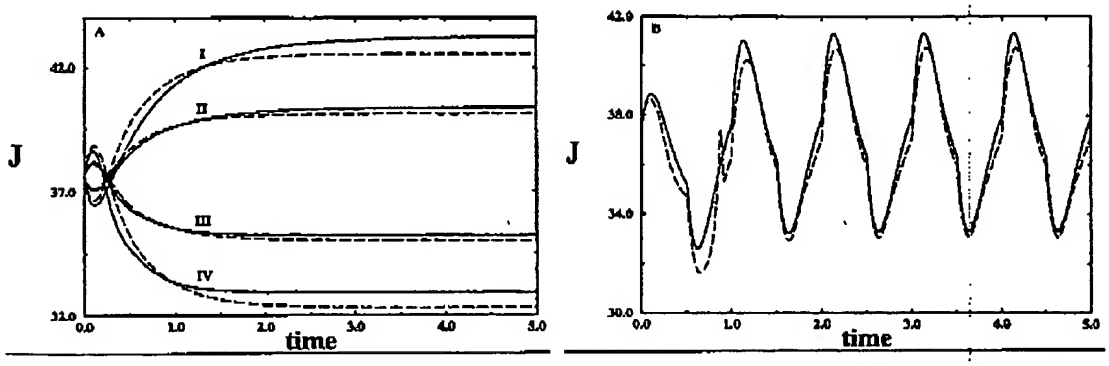

Figure 2. Dynamic responses of the ethanol specific production rate of the yeast glycolytic pathway. Solid lines correspond to the nonlinear model and dashed lines correspond to the (log)linear model. A & C. Responses to step changes of the glucose uptake. I: +20%; II: +10%; III: −10%; IV: −20% (from Hatzimanikatis and Bailey, 1997b.) B. Responses to periodic variation of the input flux with period $T = 1$ and $v_{in}(t) = 1.2v_{in,o}$ for $nT \leq t < (n + 0.5)T$ and $v_{in}(t) = 0.8v_{in,o}$ for $(n + 0.5)T \leq t < (n + 1)T$.

Optimization of Microbial Ethanol Production

The kinetic description provided by the (log)linear metabolic model has been used for identification of the changes in the amounts of the enzymes of the glycolytic pathway and of the modifications in the regulatory characteristics of those enzymes that will maximize ethanol production.

The problem of maximizing microbial ethanol production has been addressed in two different contexts: the *analysis problem* and the *synthesis problem*. In the context of the regulatory structure present in yeast, which can be modified by inactivating any of the regulatory loops, the *analysis problem* is postulated as:

> *Which of the existent regulatory loops should be inactivated, and what associated changes should be made in enzyme expression levels, in order to maximize specific ethanol production?*

In a more extended context, a *regulatory superstructure* is postulated which combines the regulatory characteristics of the glycolytic pathway in the bacterium, *E. coli* and yeast, and the *synthesis problem* considers the following questions:

> *How should the regulatory loops present in both* E. coli *and yeast be combined, and what associated changes should be made in enzyme expression levels in order to maximize specific ethanol production?*

The number of the regulatory loops in the *synthesis problem* is equal to 7 and the number of the enzymes that can be manipulated is equal to 8. Therefore, there are $2^{(7+8)} = 32{,}768$ alternative solution structures. In the *synthesis problem*, the number of the regulatory loops is equal to 10 and the number of the enzymes that can be manipulated is equal to 8, resulting in $2^{(10+8)} = 262{,}144$ alternative solution structures. The highly combinatorial nature of the problem manifests the need for a mathematical and computational method that can screen for the most promising initial strategies.

The optimization calculation for the *analysis problem* indicated that only two activation loops should be retained, and three enzymes should be overexpressed in order to maximize ethanol production, giving an increase of 100% relative to the reference state. The optimal regulatory structure and enzyme manipulation pattern is illustrated in Figure 3.

Solution of the *synthesis problem*, which begins from a pathway containing the superimposed enzymatic regulation of *E. coli* and of *S. cerevisiae* (Fig. 4) suggested the regulatory structure presented in Figure 5 and overexpression of three enzymes, with a higher overexpression level for pyruvate kinase (Pyk), that could increase ethanol production by 114%. In the optimal glycolytic pathway suggested by the calculations, phosphofructokinase (Pfk) features regulatory characteristics from *E. coli*, and Pyk is affected by only one regulatory loop, activation by FdP, common in both *E. coli* and yeast.

We could, therefore, draw the conclusion that, in an ethanol-producing *E. coli* strain, overexpression of a Pyk that

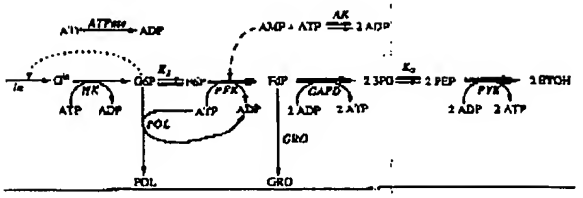

Figure 3. The optimal regulatory structure and enzyme manipulation calculated for the analysis problem for optimization of yeast ethanol production. Thick solid arrows indicate increases in enzyme levels for the respective reaction. Abbreviations for chemical species and pathway steps are the same as in Figure 1.

lacks inhibition from ATP might increase the ethanol production rate. Moreover, the optimization solution suggested that overexpression of Pfk will not increase ethanol production significantly. These conclusions are strictly qualitative because the main parameters used, i.e., substrate elasticities, are based on a kinetic model for yeast, and certain stoichiometric structures from yeast, such as glucose phosphorylation and glycerol production, which are not present in *E. coli*.

The suggested solution for the synthesis problem was compared with results from laboratory experiments with an ethanol-producing *E. coli* strain under anaerobic, nongrowing conditions. The ethanol-producing *E. coli* strain KO20 (Ohta et al., 1991), obtained from Prof. L.O. Ingram (University of Florida), was used as the host in these experiments. The strain is characterized by one chromosomal copy of the artificial *pet* operon encoding the *Zymomonas mobilis* genes for alcohol dehydrogenase II and pyruvate decarboxylase, integrated within the pyruvate formate lyase gene. Thus, under anaerobic conditions, this strain produces mainly ethanol as the catabolic product of glucose utilization. Based on the *E. coli* expression vector p*Trc99*A, (Pharmacia LKB, Uppsala, Sweden), several plasmids carrying homologous and heterologous *pfk* and *pyk* genes under the inducible *tac* promoter were constructed, allowing moderate overexpression of the corresponding enzymes (2–10 times). The enzymes expressed in combinations were the *E. coli* PfkA and PykF, and the heterologous non-allosteric, nonregulated Pfk from *Lactobacillus bulgaricus* (Branny et al., 1993; Le Bras et al., 1991) and Pyk from *Bacillus stearothermophilus* (Sakai and Ohta, 1993), which lacks any activation by FdP, but is allosterically activated by AMP and ribose-5-phosphate (Sakai et al., 1986). Details of the experimental procedure can be found elsewhere (Emmerling et al., in preparation). In all experiments reported here, cells were cultivated under anaerobic, non-growing conditions. In order to determine specific glucose uptake and ethanol production rates, the extracellular glucose and ethanol concentrations were determined at successive time points.

Specific ethanol production rate was increased by 11% in the strain overexpressing the homologous *pyk* (PYKec), and by 17% in the strain overexpressing the heterologous *pyk* from *B. stearothermophilus* (PYKbs), under conditions of low initial glucose concentration and for cells harvested

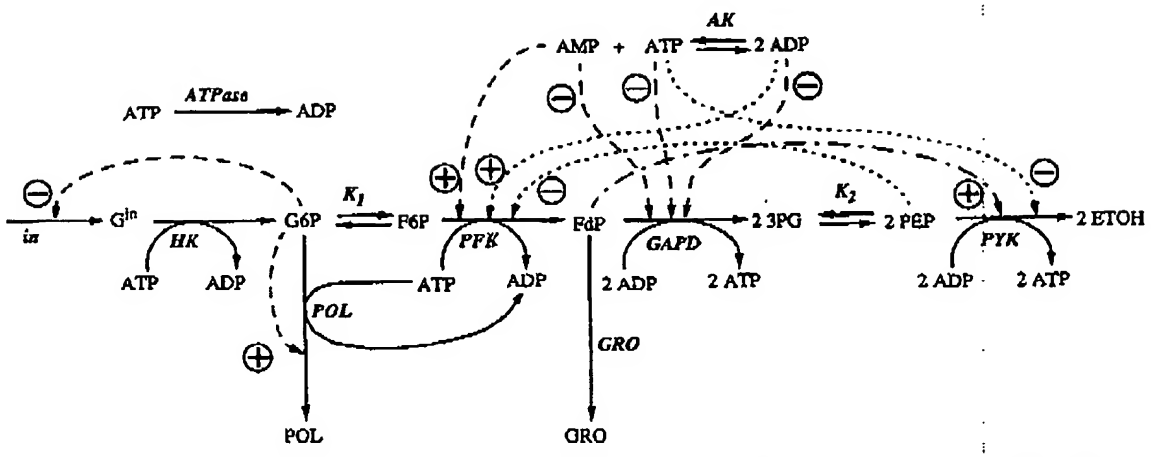

Figure 4. The glycolytic pathway with the regulatory superstructure composed by superimposing regulatory characteristics from yeast (dashed arrows) and from *E. coli* (dotted arrows). Dashed-dotted arrows indicated regulatory loops common to the two organisms. Plus signs indicate activation and minus signs indicate inhibition for the respective loops. Abbreviations for chemical species and pathway steps are the same as in Figure 1.

from mid-exponential growth phase (Fig. 6A). Moreover, specific glucose uptake rate was increased by 10% in strain PYKbs compared to the identical uptake rates of the control strain carrying only the plasmid and the PYKec strain (Fig. 6B). Under these conditions, constant glucose uptake rate and ethanol production rate were observed over the time course of the experiment (1 h) for all strains, while a decrease in the glucose uptake rate over time was observed for all strains under conditions of high initial glucose concentration, and for cells harvested from late exponential growth phase (Figs. 7A & B). However, in the PYKbs strain, glucose uptake rate was higher by 18% as compared to the control and the PYKec strains (Fig. 7B). In addition, no apparent decrease in the specific ethanol production rate was observed in the PYKbs strain, contrary to the control and the PYKec strains in which ethanol production rate decreased over time (Fig. 7A). Thus, under the conditions investigated, moderate overexpression of the heterologous *pyk* from *B. stearothermophilus* increases the glycolytic flux in *E. coli* engineered for ethanol production by 10% to 20%, and the final ethanol yield by 20% to 30%.

These presented experimental results are in qualitative agreement with the suggested solution from the synthesis problem. Strain PYKbs has a regulatory structure similar to the calculated optimal structure. However, Pyk from *B. stearothermophilus* is not activated by FdP, whereas the optimal structure suggests the desirability of this interaction.

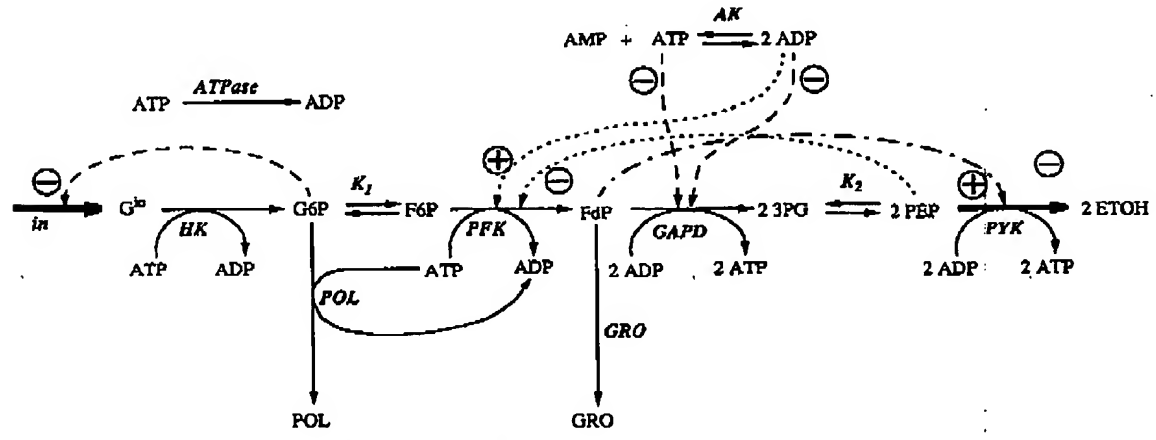

Figure 5. The optimal regulatory structure and enzyme manipulation calculated for the synthesis problem for maximizing ethanol production. Dashed-doted arrows indicate regulatory loops common to the two organisms. Plus signs indicate activation and minus signs indicate inhibition for the respective loops. Abbreviations for chemical species and pathway steps are the same as in Figure 1.

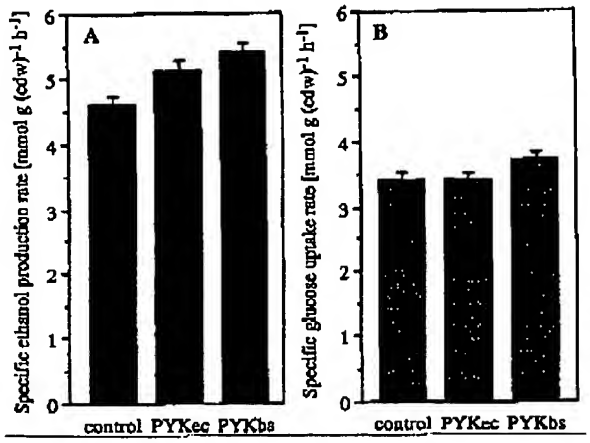

Figure 6. Specific glucose uptake (A) and ethanol production rate (B) at 0.5% initial glucose concentration of anaerobic, resting *E. coli* KO20 strains with the plasmid p*Trc99*A only (control), overexpressing *pfkF* from *E. coli* (PYKec), and overexpressing *pyk* from *B. stearothermophilus* (PYKbs). Cells were harvested from anaerobic mid-exponential growth phase cultures. The indicated values represent the mean value of at least two independent experiments, and error bars indicate the standard deviation.

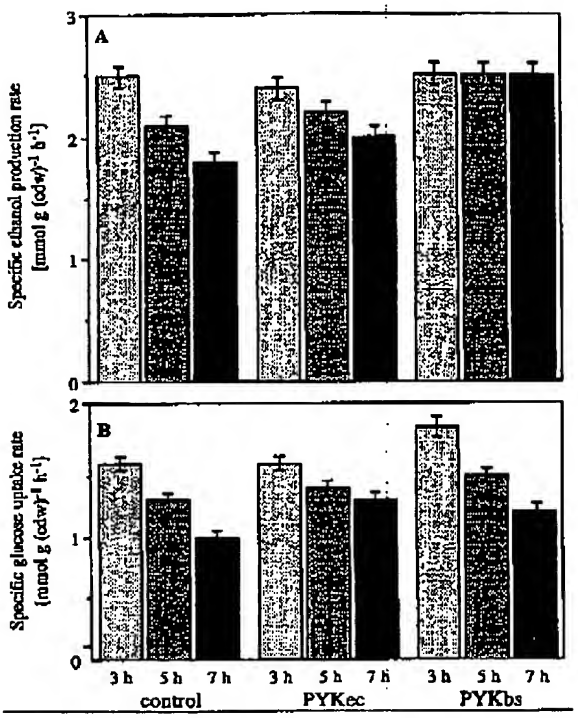

Figure 7. Specific glucose uptake (A) and ethanol production rate (B) at 2% initial glucose concentration of anaerobic, resting *E. coli* KO20 strains with the plasmid p*Trc99*A only (control), overexpressing *pfkF* from *E. coli* (PYKec), and overexpressing *pyk* from *B. stearothermophilus* (PYKbs). Different shading indicates different time points after the start of the experiment: 3 h (light grey), 5 h (grey), and 7 h (dark grey). Cells were harvested from the late exponential growth phase of aerobically grown cultures. The indicated values represent the mean value of at least two independent experiments, and error bars indicate the standard deviation.

CONCLUSIONS

A precise quantitative description of metabolism will, in most cases, be impossible, at least in the foreseeable future. However, careful mathematical modeling and methods that can take into account uncertainty in data and modeling, and incorporate biological and process constraints can provide a useful *qualitative* description of the responses of metabolism to metabolic engineering manipulations. The qualitative consistency of experimental results with trends suggested by optimization calculations using a (log)linear model indicate the utility of these mathematical tools for metabolic design.

APPENDIX

Details of the construction of the (log)linear model can be found elsewhere (Hatzimanikatis et al., 1996; Hatzimanikatis and Bailey, 1997a). Here, the structure and the parameters of the (log)linear model used for the yeast glycolysis are summarized. The nonlinear model used to construct the (log)linear model can be found elsewhere (Hatzimanikatis and Bailey, 1997a; Schlosser et al., 1994). The mass balances on the metabolites of the system may be written as:

$$\frac{d\mathbf{z}}{dt} = \mathbf{N}\varepsilon\mathbf{z} + \mathbf{N}\Pi\mathbf{q} \tag{1}$$

and the metabolic outputs can be written as:

$$\mathbf{w} = \Xi\varepsilon\mathbf{z} + \Xi\Pi\mathbf{q} \tag{2}$$

where, $\mathbf{z}$, $\mathbf{q}$, and $\mathbf{w}$, are the *logarithmic* deviations of the metabolite concentrations, the enzyme levels, and the metabolic outputs, respectively:

$$\begin{aligned} z_i &= \ln(x_i/x_{i,o}) \\ q_k &= \ln(p_k/p_{k,o}) \\ w_l &= \ln(h_l/h_{l,o}) \end{aligned} \tag{3}$$

and $\mathbf{N}$, Ξ, ε, and Π, are matrices, defined as:

$$\begin{aligned} \mathbf{N} &= \{n_{i,j} \mid n_{i,j} = \frac{\nu_{j,o}}{x_{i,o}}\left(\frac{\partial f_i}{\partial \nu_j}\right)_{x_o,p_o}\}, \\ \Xi &= \{\xi_{l,j} \mid \xi_{l,j} = \frac{\nu_{j,o}}{h_{l,o}}\left(\frac{\partial h_l}{\partial \nu_j}\right)_{x_o,p_o}\}, \\ \varepsilon &= \{\epsilon_{j,i} \mid \epsilon_{j,i} = \frac{x_{i,o}}{\nu_{j,o}}\left(\frac{\partial \nu_j}{\partial x_i}\right)_{x_o,p_o}\}, \\ \Pi &= \{\pi_{j,k} \mid \pi_{j,k} = \frac{p_{k,o}}{\nu_{j,o}}\left(\frac{\partial \nu_j}{\partial p_k}\right)_{x_o,p_o}\} \end{aligned} \tag{4}$$

The metabolite vector, $\mathbf{x}$, is defined for this model as:

$$\mathbf{x}^T = [G^{in}, F6P, FdP, PEP, ATP]$$

where the italic symbols indicate the intracellular concentrations of the corresponding metabolites.

The reaction vector, v, is here:

$$v^T = [v_{in}, v_{HK}, v_{POL}, v_{PFK}, v_{GRO}, v_{GAPD}, v_{PYK}, v_{ATPase}]$$

The elasticity matrix ε can be written as a sum of two matrices:

$$\varepsilon = \varepsilon^s + \varepsilon^r \quad (5)$$

where the elements in matrix ε^s correspond to the substrate elasticities of the enzymes, that is, the sensitivities of enzyme activities with respect to concentrations of their substrates, and the elements of matrix ε^r correspond to the regulatory elasticities of the enzymes, that is, the sensitivities of enzyme activities with respect to concentrations of regulatory metabolites.

Changes of the elements in matrix ε^r from non-zero values to zero, or vice-versa, correspond to modifications in the regulatory structure of the system.

The stable steady state used as the reference state from which the linear model is constructed is:

$$x_0^T = [0.080200, 2.56666, 9.29469, 0.015318, 0.819456]$$

and the values for the corresponding steady state rates are:

$$v_0^T = [27.366669\ \ 27.366669\ \ 4.374271\ \ 22.992396\ \ 18.804933\ \ 4.187464\ \ 37.609866\ \ 20.486395]$$

For the matrix N in Equation (1) we have:

$$N = \begin{bmatrix} 341.230 & -341.230 & 0 & 0 \\ 0 & 10.6624 & -1.70426 & -8.95809 \\ 0 & 0 & 0 & 2.47371 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -33.3961 & -5.33802 & -28.0581 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -2.02319 & -0.450522 & 0 & 0 \\ 0 & 546.738 & -2455.27 & 0 \\ 0 & 10.2201 & 45.8961 & -24.0 \end{bmatrix}$$

The elasticity matrices at the reference steady state are:

$$\varepsilon^s = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.55174 & 0 & 0 & 0 & 0 \\ 0 & 0.20134 & 0 & 0 & 0.05252 \\ 0 & 0.41708 & 0 & 0 & 0.03270 \\ 0 & 0 & 0.72084 & 0 & 0 \\ 0 & 0 & 0.054201 & 0 & -0.06921 \\ 0 & 0 & 0 & 0.35162 & 0.07157 \\ 0 & 0 & 0 & 0 & 1.0 \end{bmatrix}$$

and

$$\varepsilon^r = \begin{bmatrix} 0 & -0.04689 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0.93425 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.230352 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.16094 \\ 0 & 0 & 0.04261 & 0 & 0.02834 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We consider the quantities of the enzymes catalyzing the first six reactions as the manipulated variables. Therefore, we can write for the manipulated variables the vector:

$$p^T = [V_{m,in}, V_{m,HK}, V_{m,POL}, V_{m,PFK}, V_{m,GRO}, V_{m,GAPD}, V_{m,PYK}, V_{m,ATPase}]$$

and for the matrix Π we have:

$$\Pi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Ethanol production has been considered to be proportional to the rate of the reaction catalyzed by pyruvate kinase. Therefore, the parameters in Equation (2) that correspond to v_{PYK} are:

$$\Xi_{PYK} = 1$$

and ε^s and ε^r as given above.

The synthesis problem was considered as in Hatzimanikatis et al. (1996).

References

Bailey, J. E. 1991. Toward a science of metabolic engineering Science 252: 1668–1675.

Boles, E., Heinisch, J., Zimmermann, F. K. 1993. Different signals control the activation of glycolysis in the yeast *Saccharomyces cerevisiae*. Yeast 9: 761–770.

Branny, P., de la Torre, F., Garel, J.-R. 1993. Cloning, sequencing, and expression in *Escherichia coli* of the gene coding for phosphofructokinase in *Lactobacillus bulgaricus*. J. Bacteriol. 175: 5344–5349.

Cornish-Bowden, A., Cardenas, M. L. (eds.) 1990. Control of metabolic processes. NATO ASI Series A: Lifesciences Vol. 190. Plenum Press, New York.

Fell, D. A. 1992. Metabolic control analysis–A survey of its theoretical and experimental development. Biochem. J. 152: 313–330.

Gallazo, J. L., Bailey, J. E. 1990. Fermentation pathway kinetics and metabolic flux control in suspended and immobilized *Saccharomyces cerevisiae*. Enzyme Microb. Technol. 12: 162–172.

Hatzimanikatis, V., Floudas, C. A., Bailey, J. E. 1996. Analysis and design of metabolic reaction networks via mixed-integer linear optimization. AIChE J. 42: 1277–1292.

Hatzimanikatis, V., Bailey, J. E. 1997a. Effects of spatiotemporal variations on metabolic control: Approximate analysis using (log)linear kinetic models. Biotech. Bioeng. 54: 91–104.

Hatzimanikatis, V., Bailey, J. E. 1997b. Studies on glycolysis: I. Multiple steady states in bacterial glycolysis. Chem. Engng Sci. 52: 2579–2588.

Heinrich, R., Rapoport, S. M., Rapoport, T. A. 1977. Metabolic regulation and mathematical models. Prog. Biophys. Mol. Biol. 32: 1–82.

Joshi, A., Palsson, B. O. 1989. Metabolic dynamics in the human red cell. Part I—A comprehensive kinetic model. J. Theor. Biol. 141: 515–528.

Ingram, L. O., Doran, J. B. 1995. Conversion of cellulosic materials to ethanol. FEMS Microbiol. Rev. 16: 235–241.

Kacser, H., Burns, J. A. 1973. The control of flux. Symp. Soc. Exp. Biol. 27: 65–104.

Katsumata, R., Ikeda, M. 1993. Hyperproduction of tryptophan in *Corynebacterium glutamicum* by pathway engineering. Bio/Technol 11: 921–925.

Le Bras, G., Deville-Bonne, D., Garel, J.-R. 1991. Purification and properties of the phosphofructokinase from *Lactobacillus bulgaricus*: A non-allosteric analog of the enzyme from *Escherichia coli*. Eur. J. Biochem. 198: 683–687.

Ohta, K., Beall, D. S., Mejia, J. P., Shanmugam, K. T., Ingram, L. O. 1991. Genetic improvement of *Escherichia coli* for ethanol production: Chromosomal integration of *Zymomonas mobilis* genes encoding pyruvate decarboxylase and alcohol dehydrogenase II. Appl. Environ. Microbiol. 57: 893–900.

Sakai, H., Ohta, T. 1993. Molecular cloning and nucleotide sequence of the gene for the pyruvate kinase of *Bacillus stearothermophilus* and the production of the enzyme in *Escherichia coli*. Eur. J. Biochem. 211: 851–859.

Sakai, H., Suzuki, K., Imahori, K. 1986. Purification and properties of pyruvate kinase from *Bacillus stearothermophilus*. J. Biochem. 99: 157–1167.

Schaaff, I., Heinisch, J., Zimmermann, F. K. Overproduction of glycolytic enzymes in yeast. Yeast 5: 285–290.

Schlosser, P. M., Holcomb, T., Bailey, J. E. 1993. Determining metabolic sensitivity coefficients directly from experimental data. Biotechnol. Bioeng. 41: 1027–1038.

Schlosser, P. M., Riedy, G., Bailey, J. E. 1994. Ethanol production in baker's yeast: Application of experimental perturbation techniques for model development and resultant changes in flux control analysis. Biotechnol. Prog. 10: 141–154.